STATE UNIVERSITY OF NEW YORK
COLLEGE OF TECHNOLOGY
CANTON, NEW YORK

MASTER SYLLABUS

CYBR 315 – DATA MINING AND MACHINE LEARNING

Created by: Paweena Waidelich
Updated by:
A. **TITLE:** Data Mining and Machine Learning

B. **COURSE NUMBER:** CYBR 315

C. **CREDIT HOURS (Hours of Lecture, Laboratory, Recitation, Tutorial, Activity):**

# Credit Hours: 3
# Lecture Hours per Week: 3
# Lab Hours per Week:
Other per Week:

**Course Length (# of Weeks):** 15 weeks (or 7 weeks)

D. **WRITING INTENSIVE COURSE:** n/a

E. **GER CATEGORY:** n/a

F. **SEMESTER(S) OFFERED:** Fall and Spring

G. **COURSE DESCRIPTION:** This course provides the concepts and techniques in processing gathered data or information, which will be used in various applications. Specifically, it explains data mining and the tools used in discovering knowledge from the collected data. This course focuses on the feasibility, usefulness, effectiveness, and scalability of techniques of large data sets. After describing data mining, this course explains the methods of knowing, preprocessing, processing, and warehousing data. Then, the methods involved in mining frequent patterns, associations, and correlations for large data sets are described. The course details the methods for data classification and introduces the concepts and methods for data clustering. Finally, it discusses the outlier detection and the trends, applications, and research frontiers in data mining.

H. **PRE-REQUISITES/CO-REQUISITES:**

   a. Pre-requisite(s): None.

I. **STUDENT LEARNING OUTCOMES:**

<table>
<thead>
<tr>
<th>Course Student Learning Outcome [SLO]</th>
<th>GER</th>
<th>ISLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Examine the field of data science and familiarize with basics of data mining and machine learning.</td>
<td>2. Critical Thinking [CA]</td>
<td></td>
</tr>
<tr>
<td>b. Identify the process of collecting the data, pre-process and preparing it for data analysis.</td>
<td>2. Critical Thinking [CA]</td>
<td></td>
</tr>
<tr>
<td>c. Apply data classification and clustering methods on wide range of data to practice and understand how the methods work.</td>
<td>2. Critical Thinking [PS]</td>
<td></td>
</tr>
<tr>
<td>ISLO #</td>
<td>ISLO &amp; Subsets</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------</td>
<td></td>
</tr>
</tbody>
</table>
| 1      | Communication Skills  
Oral [O], Written [W] |
| 2      | Critical Thinking  
Critical Analysis [CA], Inquiry & Analysis [IA], Problem Solving [PS] |
| 3      | Foundational Skills  
Information Management [IM], Quantitative Lit./Reasoning [QTR] |
| 4      | Social Responsibility  
Ethical Reasoning [ER], Global Learning [GL], Intercultural Knowledge [IK], Teamwork [T] |
| 5      | Industry, Professional, Discipline Specific Knowledge and Skills |

**J. APPLIED LEARNING COMPONENT:**  
Yes______ No ____

If Yes, select one or more of the following categories:

- Classroom/Lab___  
- Internship___  
- Clinicial Practicum___  
- Practicum___  
- Service Learning___  
- Community Service___

- Civic Engagement___  
- Creative Works/Senior Project___  
- Research___  
- Entrepreneurship___  

(program, class, project)
I. Introduction
   A. What is Data Mining and what data can be mined?
   B. What patterns can be detected
II. Getting to know your data
   A. Data Objects and Attribute Types
   B. Basic Statistical Descriptions of Data
   C. Data Visualization
III. Data Preprocessing
   A. Data Cleaning
   B. Data Integration
   C. Data Reduction
   D. Data Discretization
IV. Mining Frequent Patterns, Associations, and Correlations
   A. Basic Concepts
   B. Frequent Itemset Mining Methods
   C. Which Patterns Are Interesting?—Pattern Evaluation Methods
V. Classification
   A. Decision Tree Induction
   B. Bayes Classification Methods
   C. Rule-Based Classification
   D. Model Evaluation and Selection
VI. Cluster Analysis: Basic Concepts and Methods
   A. Cluster Analysis
   B. Partitioning Methods
   C. Hierarchical Methods
   D. Density-Based Methods
   E. Evaluation of Clustering
VII. Outlier Detection
   A. Outliers and Outlier Analysis
B. Outlier Detection Methods
C. Statistical Approaches
D. Proximity-Based Approaches

Q. LABORATORY OUTLINE:
n/a